Course: ELEC/COMP 526
Term: Spring 2020
Room: Building DH 1042
Class: T. Th, 9:25-10:40

INSTRUCTOR CONTACT INFORMATION

Instructor: Peter Varman
Office: DH 2022
Tel: 713-348-3990
Email: pjv@rice.edu
Webpage: www.ece.rice.edu/~pjv
Office Hours: T, TH 2:00 – 4:00 pm or by Appointment

COURSE OBJECTIVES AND LEARNING OUTCOMES

The goal of the course is to provide students with an advanced understanding of the architecture of modern computer systems designed for high performance.

Following the course the students will be able to:

(a) describe the handling of data and control dependencies in high-performance uniprocessors
(b) design a microarchitecture for pipelined and dynamically scheduled out-of-order processors
(c) analyze the behavior of high-performance memory systems including non-blocking and lockup-free caches
(d) describe the problems of cache coherence in multiprocessor systems
(e) implement and analyze the behavior of snooping and directory-based cache coherence protocols
(f) describe the problems of memory ordering and strong and weak consistency in multiprocessors
(g) describe the implementation of virtual memory and its interaction with the cache hierarchy
(h) describe architectural support for synchronization including transactional memory
(i) describe the architecture of modern storage and network systems for datacenter computing
(j) have familiarity with elements of datacenter computing and system virtualization concepts

RECOMMENDED TEXTS AND MATERIALS

- **Fundamentals of Parallel Multicore Architecture**, **Yan Solihin**
- Parallel Computer Organization and Design, M. Dubois, M. Annavaram, P. Stenström
- Computer Architecture: A Quantitative Approach, (5th edition), J. Hennessy and D. Patterson
- Survey and Research Papers (Will be posted as needed)

PROJECTS AND EXAM

Homework Sets and Projects (5-6)	(Assignment-specific weights)	60
Semester Tests	(2 Evening or in-class tests)	25
Individual Project		15

Projects will involve C programming to implement simulation models, running simulation experiments, and analyzing the results.

LATE POLICIES

All homework assignments and quizzes are due by the deadlines. There is a 5% for up to two days beyond the deadline, and a 10% penalty per day for the next three.

RICE HONOR CODE

In this course, all students will be held to the standards of the Rice Honor Code, a code that you pledged to honor when you matriculated at this institution. If you are unfamiliar with the details of this code and how it is administered, you should consult the Honor System Handbook at http://honor.rice.edu/honor-system-handbook/. This handbook outlines the University's expectations for the integrity of your academic work, the procedures for resolving alleged violations of those expectations, and the rights and responsibilities of students and faculty members throughout the process.

- All assignments and tests are assumed to represent the individual efforts of a student and are given under the (letter and spirit) of the Rice University Honor Code system.
- Do not consult homework sets, quizzes, solutions or tests from past versions of this class.
- Discussion with other students and teaching staff on assignments is encouraged.
- All submitted work must be done independently by each individual (or lab group if applicable).

DISABILITY SUPPORT SERVICES

If you have a documented disability or other condition that may affect academic performance you should: 1) make sure this documentation is on file with Disability Support Services (Allen Center, Room 111 / adarice@rice.edu / x5841) to determine the accommodations you need; and 2) talk with me to discuss your accommodation needs.

SYLLABUS CHANGE POLICY

This syllabus is only a guide for the course and is subject to change with advance notice.

Topics

1. **Processor**
    - Instruction Level Parallelism
        - Simple static pipelining
        - Dynamically-scheduled pipeline processors
        - Superscalar, VLIW
    - Data Parallelism
        - Identical operations on elements of large arrays (vectors)
            - Vector and Stream Processors, GGPUs
            -
    - Thread Level Parallelism
        - Application structured as multiple cooperating threads
            - Multithreaded Processors
            - Multi Cores
            - NUMA and multiprocessor processors
            - Clusters

2. **Memory**
    - Uniprocessor Caches
        - Organizations and Addressing
        - Lockup Free Caches
        - Hierarchical Cache Organization
    - Multiprocessor Caches
        - Cache Coherence
        - Snooping and Directory Protocols
        - MSI, MSEI, MSEOI protocols
    - Memory Orderings and Consistency
    - Virtual Memory
        - Demand Paging
        - TLB and Page Walks
        - Virtually Indexed and Physically Tagged Caches

3. **Synchronization**
    - Spin Locks
    - Sleep Locks
    - Lock-Free Synchronization
    - Transactional Memory

4. **Interconnect and Networking**
    - Bandwidth, Latency, Topology
    - Routing
    - Scalable Networks : Close, Butterfly, Fat Trees
    - RDMA (InfiniBand, RoCE)

5. **Storage Systems**
    - Storage Architecture: DAS, SAN, NAS, Cluster
    - Hybrid Storage Arrays: Disks, SSDs, NVM

- Distributed Storage: Reliability, RAID, Erasure Codes
- Byte Addressable NVRAM
-

6. **System Resource Virtualization** (Software Defined Everything (SDX)
    - Sharing vs Isolation
    - Resource Virtualization Techniques
    - Resource Scheduling